Bellwork

Check whether each of the conditions is met for calculating a confidence interval for the population proportion p.

High tuition costs Glenn wonders what proportion of the students at his school think that tuition is too high. He interviews an SRS of 50 of the 2400 students at his college. Thirty-eight of those interviewed think tuition is too high.

Random : yes, SRS

Independent : $50 \le .1(2700)$ yes

Normal : $50(.76) \ge 10$ & $50(1-.76) \ge 10$

yes

Dec 1-10:43 AM

24.

$Z^* \sqrt{\dfrac{\hat{p}(1-\hat{p})}{n}} \le ME$ invNORM (.05,0,1)

$.05 \quad .9 \quad .05$

a) $1.645 \sqrt{\left(\dfrac{.75(1-.75)}{n}\right)} \le \left(\dfrac{.04}{1.645}\right)^2$ $z = -1.645 = Z^*$

$Z^* = 1.645$

$\dfrac{.75(1-.75)}{n} \le .00059 \cdot n$

$\dfrac{.75(1-.75)}{.00059} \le \dfrac{.00059n}{.00059}$

b) 424 people $317.79 \le n$

$\boxed{318 \text{ people}}$

Jan 15-7:54 AM

$$\hat{p} \pm z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

MOE

26.

a) $z^* \sqrt{\dfrac{\hat{p}(1-\hat{p})}{n}} = ME$

$z^* \sqrt{\dfrac{.64(1-.64)}{1028}} = .03$

$z^* \dfrac{.04197}{.04197} = \dfrac{.03}{.04197}$　　$z^* = 2.0039$

Jan 15-8:07 AM

8.3
Estimating a
Population Mean
(Day 1)

vocab

examples

Dec 1-10:43 AM

Objectives

- Construct and interpret a confidence interval for a population mean

- Determine the sample size required to obtain a level *C* confidence interval for a population mean with a specified margin of error

- Carry out the 4 steps in constructing a confidence interval for a population mean

- Determine sample statistics from a confidence interval

- Understand why each of the three inference conditions - Random, Normal, and Independent - is important

Dec 1-10:43 AM

One-Sample z Interval for a Population Mean

Draw an SRS of size *n* from a population having unknown mean μ and known standard deviation σ. As long as the Normal and Independent conditions are met, a level *C* confidence interval for μ is:

$$\overline{x} \pm z^* \frac{\sigma}{\sqrt{n}}$$

The critical value z* is found from the standard Normal distribution

For means, the distribution is normal when n ≥ 30, by CLT

Dec 1-10:43 AM

Choosing Sample Size for a Desired Margin of Error when estimating μ

To determine the sample size *n* that will yield a level C confidence interval for a population mean with a specified margin of error *ME:*

- get a reasonable value for the population standard deviation σ from an earlier or pilot study

- find the critical value z* from a standard Normal curve for confidence level *C*

- Set the expression for the margin of error to be less than or equal to *ME* and solve for *n*:

$$z^* \frac{\sigma}{\sqrt{n}} \le ME$$

Dec 1-10:43 AM

How Many Mokeys?

Researchers would like to estimate the mean cholesterol level μ of a particular variety of monkey that is often used in laboratory experiments. They would like their estimate to be within 1 milligram per deciliter (mg/dl) of the true value of their estimate at a 95% confidence level. A previous study involving this variety of monkey suggests that the standard deviation of cholesterol level is 5 mg/dl.

Obtaining monkeys is time-consuming and expensive, so the researchers want to know the minimum number of monkeys they will need to generate a satisfactory estimate.

$$1.96 \cdot \frac{5}{\sqrt{n}} \le 1$$

$$\frac{9.8}{\sqrt{n}} \le 1 \sqrt{n}$$

$$(9.8)^2 \le (\sqrt{n})^2$$

$$96.04 \le n$$

97 monkeys

.95 .025 invNorm.

Dec 1-10:43 AM

8.3
Estimating a
Population Mean
(Day 2)

vocab

examples

Dec 1-10:43 AM

Calculator BINGO!

A farmer wants to estimate the mean weight (in grams) of all tomatoes grown on his farm. To do so, he will select a random sample of 4 tomatoes, calculate the mean weight (in grams), and use the sample mean to create a 99% confidence interval for the population mean. Suppose that the weights of tomatoes on his farm are approximately Normally distributed with a mean of 100 grams and a standard deviation of 40 grams.

Step 1. Use your calculator to simulate taking an SRS of size 4 from this population and creating a one-sample *z* interval

Enter the command shown below and press ENTER. Check to see if the resulting interval captures = 100. If it does not, shout "BINGO!"

    randNorm(100,40,4) --> L1:Z Interval 40, mean(L1), 4, 99

Keep pressing ENTER!

Dec 1-10:43 AM

Calculator BINGO!

The method in Step 1 works well if we know the population standard deviation σ. That's rarely the case in real life. What happens if we use the sample standard deviation in place of σ when calculating a confidence interval for the population mean?

Step 2. Use your calculator to simulate taking an SRS of size 4 from this population and creating a "modified" one-sample z interval

Enter the command shown below and press ENTER. Check to see if the resulting interval captures = 100. If it does not, shout "BINGO!"

        randNorm(100,40,4)-->L1:Z Interval stdDev(L1), mean(L1), 4, 99

Keep pressing ENTER!

Dec 1-10:43 AM

Calculator BINGO!

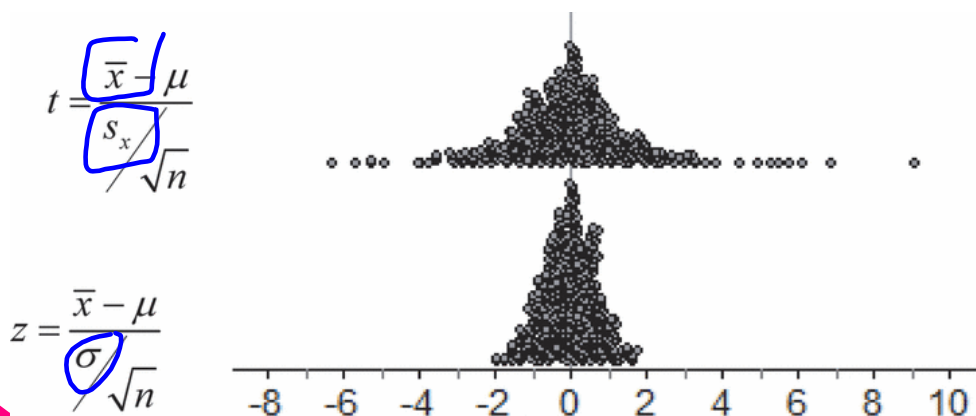How do the two methods compare? Did one method do a better job of capturing the true mean? Why?

Dec 1-10:43 AM

t distribution

Shape: *Symmetric, unimodal, but not quite Normal (heavier tails)*

Center: 0 - still a standardized scored

Spread: Greater than a Normal distribution,
        Gets closer to Normal as degrees of freedom increase

$$t = \frac{\overline{x} - \mu}{s_x / \sqrt{n}}$$

$$z = \frac{\overline{x} - \mu}{\sigma / \sqrt{n}}$$

-8 -6 -4 -2 0 2 4 6 8 10

Dec 1-10:43 AM

t-distribution: Degrees of Freedom

Drawn an SRS of size *n* from a large population that has a Normal distribution with mean μ and standard deviation σ. The statistic:

$$t = \frac{\overline{x} - \mu}{s_x / \sqrt{n}}$$

has the t distribution with degrees of freedom df = n-1. This statistic will have approximately a t_{n-1} distribution as long as the sampling distribution of $\overline{x}$ is close to Normal.

Dec 1-10:43 AM

7

t-distribution: Degrees of Freedom

t distributions have more area in the tails than the standard Normal distribution.

- - - t, 2 degrees of freedom
- ⋯⋯ t, 9 degrees of freedom
- —— Standard Normal

0
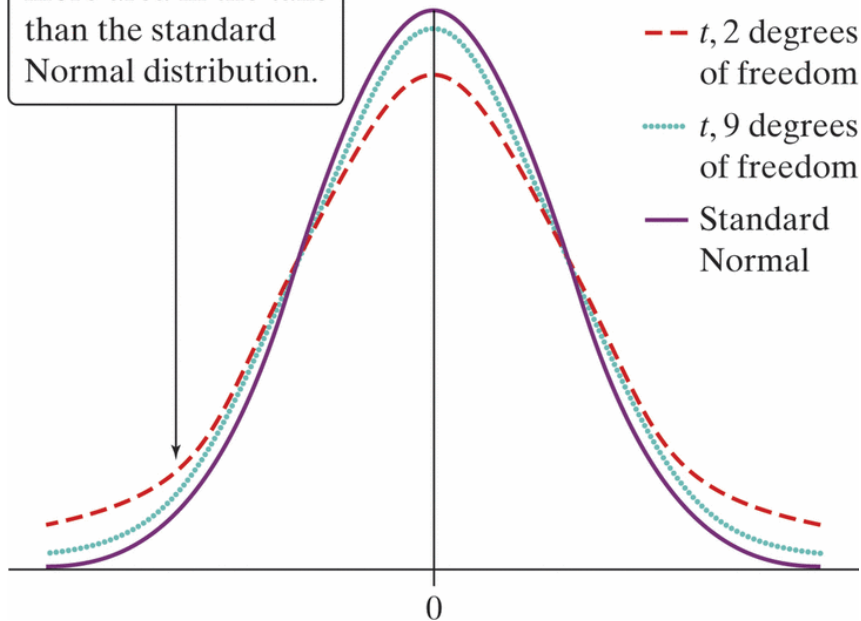
Dec 1-10:43 AM

t-distribution: Degrees of Freedom

df	Upper-tail probability p			
	.05	.025	.02	.01
10	1.812	2.228	2.359	2.764
11	1.796	2.201	2.328	2.718
12	1.782	2.179	2.303	2.681
z^*	1.645	1.960	2.054	2.326
	90%	95%	96%	98%
	Confidence level C			

Dec 1-10:43 AM

Example

(a) Suppose you wanted to construct a 90% confidence interval for the mean of a Normal population based on an SRS of size 10. What critical value t^* should you use?

$$df = 10 - 1 = 9 \qquad \boxed{t^* = 1.833}$$

(b) What if you wanted to construct a 99% confidence interval for using a sample of size 75?

$$t^* = 2.66 \qquad df = 60$$

on calculator

Dec 1-10:43 AM

Example

(a) Suppose you wanted to construct a 90% confidence interval for the mean of a Normal population based on an SRS of size 10. What critical value t^* should you use?

(b) What if you wanted to construct a 99% confidence interval for using a sample of size 75?

*df = 75-1 = 74 *use 60**
$t^* = 2.66$

on calculator

Dec 1-10:43 AM

9

Standard error of the sample mean

The standard error of the sample mean, $\overline{x}$ is: $\dfrac{s_x}{\sqrt{n}}$

$\sigma_{\overline{x}} = \dfrac{\sigma}{\sqrt{n}}$

where s_x is the sample standard deviation.

notation: $SE_{\overline{x}}$

It describes how far $\overline{x}$ will be from μ on average in repeated SRSs of size n.

formula sheet?

Dec 1-10:43 AM

The One-Sample t Interval for a Population Mean

Choose an SRS of size *n* from a population having unknown mean μ. A level C confidence interval for μ is

$$\overline{x} \pm t^* \, \dfrac{s_x}{\sqrt{n}}$$

where t* is the critical value for the t_{n-1} distribution.

Use this interval only when (1) the population distribution is Normal or the sample size is large (n ≥ 30), and (2) the population is at least 10 times as large as the sample (10% condition)

formula sheet?   No :)

Dec 1-10:43 AM

Conditions for Inference about a Population Mean

Random: the data come from a random sample of size *n* from the population of interest or a randomized experiment. This condition is very important!

Normal: The population has a Normal distribution or the sample size is large (n≥ 30)

Independent: 10% condition is met!

Dec 1-10:43 AM

8.3
Estimating a
Population Mean
(Day 3)

vocab

examples

Dec 1-10:43 AM

How Much Homework?

The principal at a large high school claims that students at his school spend at least 10 hours per week doing homework, on average. To investigate this claim, an AP Statistics class selected a random sample of 250 students from their school and asked them how long they spent doing homework during the last week. The sample mean was 10.2 hours and the sample standard deviation was 4.2 hours.

Construct and interpret a 95% confidence interval for the mean time that all students at this school spent doing homework in the last week. Use the 4 step process.

Dec 1-10:43 AM

How Much Homework?

The principal at a large high school claims that students at his school spend at least 10 hours per week doing homework, on average. To investigate this claim, an AP Statistics class selected a random sample of 250 students from their school and asked them how long they spent doing homework during the last week. The sample mean was 10.2 hours and the sample standard deviation was 4.2 hours.

Construct and interpret a 95% confidence interval for the mean time that all students at this school spent doing homework in the last week. Use the 4 step process.

STATE: We want to estimate the mean time that all students at the school spend doing homework in the last week at a 95% confidence level

Dec 1-10:43 AM

How Much Homework?

The principal at a large high school claims that students at his school spend at least 10 hours per week doing homework, on average. To investigate this claim, an AP Statistics class selected a random sample of 250 students from their school and asked them how long they spent doing homework during the last week. The sample mean was 10.2 hours and the sample standard deviation was 4.2 hours. Construct and interpret a 95% confidence interval for the mean time that all students at this school spent doing homework in the last week. Use the 4 step process.

PLAN: Random: yes

Independent: 250 ≤ .1N Yes, assuming there are at least 2500 students at the school

Normal: yes, by CLT since n ≥ 30

We will use a one sample t-interval for μ

Dec 1-10:43 AM

How Much Homework?

The principal at a large high school claims that students at his school spend at least 10 hours per week doing homework, on average. To investigate this claim, an AP Statistics class selected a random sample of 250 students from their school and asked them how long they spent doing homework during the last week. The sample mean was 10.2 hours and the sample standard deviation was 4.2 hours. Construct and interpret a 95% confidence interval for the mean time that all students at this school spent doing homework in the last week. Use the 4 step process.

$df = 250 - 1 = 249$

$\overline{X} \pm t^* \left(\dfrac{S_x}{\sqrt{n}} \right)$

DO: $10.2 \pm 1.984 (4.2/\sqrt{250}) = 10.2 \pm .5270 = (9.63, 10.77)$

CONCLUDE: We are 95% confident that the true mean time that all students at the school spent doing homework in the last week is between 9.63 and 10.77 hours.

Dec 1-10:43 AM

Calculator...

Dec 1-10:43 AM

Normal Condition

If the sample size is small (n < 30), and we do not know the shape of the population:

- Graph the sample data

- If there are no outliers or strong skewness, it is plausible that the sample came from a Normal population (so using t is OK)

-If there are outliers or strong skewness, then don't use t when n < 30.

Dec 1-10:43 AM

AP TIP!

If the sample size is small, you must graph the sample data and comment on the Normal condition.

You must refer to the Normality of the population.

Dec 1-10:43 AM

Using One-Sample t Procedures: The Normal Condition

- **Sample size less than 15**: Use t procedures if the data appear to be close to Normal (roughly symmetric, one peak, no outliers). If the data are clearly skewed or if outliers are present, do not use t.

Dec 1-10:43 AM

Robust Procedures

An inference procedure is called robust if the probability calculations involved in the procedure remain fairly accurate when a condition for using the procedure is violated.

Dec 1-10:43 AM

Can you spare a square?

As part of their final project in AP Statistics, Christina and Rachel randomly selected 18 rolls of a generic brand of toilet paper to measure how well this brand could absorb water. To do this, they poured 1/4 cup of water onto a hard surface and counted how many squares it took to completely absorb the water. Here are the results from their 18 rolls:

| 29 | 20 | 25 | 29 | 21 | 24 | 27 | 25 | 24 |
| 29 | 24 | 27 | 28 | 21 | 25 | 26 | 22 | 23 |

Construct and interpret a 99% confidence interval for μ the mean number of squares of generic toilet paper needed to absorb 1/4 cup of water.

Dec 1-10:43 AM

Can you spare a square?

As part of their final project in AP Statistics, Christina and Rachel randomly selected 18 rolls of a generic brand of toilet paper to measure how well this brand could absorb water. To do this, they poured 1/4 cup of water onto a hard surface and counted how many squares it took to completely absorb the water. Here are the results from their 18 rolls:

29	20	25	29	21	24	27	25	24
29	24	27	28	21	25	26	22	23

Construct and interpret a 99% confidence interval for μ the mean number of squares of generic toilet paper needed to absorb 1/4 cup of water.

STATE: We want to estimate the population mean, μ, number of squares of generic toilet paper needed to absorb 1/4 cup water, at a 99% confidence level.

Dec 1-10:43 AM

Can you spare a square?

PLAN: Random - yes
Independent - yes, assuming $18 \leq .1N$, 180
ROlls of TP
Normal - since $n \not\geq 30$, we plot our data

we continue with one sample
t-interval for μ, since there
are no outliers or strong skew. The
procedure is Robust.

Dec 1-10:43 AM

Can you spare a square?

As part of their final project in AP Statistics, Christina and Rachel randomly selected 18 rolls of a generic brand of toilet paper to measure how well this brand could absorb water. To do this, they poured 1/4 cup of water onto a hard surface and counted how many squares it took to completely absorb the water. Here are the results from their 18 rolls:

29 20 25 29 21 24 27 25 24
29 24 27 28 21 25 26 22 23

Construct and interpret a 99% confidence interval for μ the mean number of squares of generic toilet paper needed to absorb 1/4 cup of water.

PLAN: ***Random*** *- yes.* ***Independent*** *- assuming that there are more than 180 brands of TP.* ***Normal*** *- since the sample size is smaller than 30, we plot the data. There are no outliers & no strong skew so we will use a one sample t-interval for μ. Note: This will be robust.*

Dec 1-10:43 AM

Can you spare a square?

As part of their final project in AP Statistics, Christina and Rachel randomly selected 18 rolls of a generic brand of toilet paper to measure how well this brand could absorb water. To do this, they poured 1/4 cup of water onto a hard surface and counted how many squares it took to completely absorb the water. Here are the results from their 18 rolls:

29 20 25 29 21 24 27 25 24
29 24 27 28 21 25 26 22 23

Construct and interpret a 99% confidence interval for μ the mean number of squares of generic toilet paper needed to absorb 1/4 cup of water.

DO: TInterval (on calculator) Inpt: Data, List: L1, Freq: 1, C-Level: .99
   $\overline{x}$ *=24.944 s_x = 2.8589*
   (22.991, 26.897)

Dec 1-10:43 AM

Can you spare a square?

As part of their final project in AP Statistics, Christina and Rachel randomly selected 18 rolls of a generic brand of toilet paper to measure how well this brand could absorb water. To do this, they poured 1/4 cup of water onto a hard surface and counted how many squares it took to completely absorb the water. Here are the results from their 18 rolls:

29	20	25	29	21	24	27	25	24
29	24	27	28	21	25	26	22	23

Construct and interpret a 99% confidence interval for μ the mean number of squares of generic toilet paper needed to absorb 1/4 cup of water.

Conclude: We are 99% confident that the interval between 22.991 and 26.897 squares of toilet paper contains the true mean number of squares that it takes to absorb 1/4 cup water.

Dec 1-10:43 AM

Homework

8.3 Assignment
Due Tuesday (no school Monday)

Dec 1-10:43 AM